

УДК 62.506.2

В. А. ЛОВИЦКИЙ

СИСТЕМА ФОРМИРОВАНИЯ ПОНЯТИИ ДЛЯ ОБЪЕКТОВ, ЗАДАНЫХ НЕЯВНЫМ НАБОРОМ ПРИЗНАКОВ

Проблеме формирования понятий с помощью ЭЦВМ в последние годы уделяется значительное внимание [1]. Интерес к ней особенно возрос после того, как возникла необходимость при построении систем с искусственным интеллектом формировать у нее представления «здорового смысла» об окружающей среде. В настоящей работе описываются принципы построения системы формирования понятий (СФП), ориентированной на работу с объектами, заданными неявным набором признаков. Все известные СФП [1] формируют понятия в процессе анализа объектов, заданных явным набором признаков. Это значит, что объекты представлены в виде последовательностей значений признаков, у которых за каждым местом закреплен один (общий для всех объектов) признак. Более строго данный вид объектов будем описывать следующим образом. Пусть задано конечное множество признаков $\Gamma = \{a_1, a_2, \dots, a_m\}$, каждый из которых может принимать конечное число различных значений. J -е значение a_i -признака обозначим через $mg_j(a_i)$. Тогда лю-

бой объект, заданный явным набором признаков, можно представить как конечную последовательность значений признаков:

$$mng_{i_1}(a_{i_1}) \rightarrow mng_{i_2}(a_{i_2}) \rightarrow \dots \rightarrow mng_{i_k}(a_{i_k}),$$

где символ \rightarrow читается как „за ... следует ...“;

$$a_{i_\nu} \neq a_{i_\mu} \text{ при } \nu \neq \mu (a_{i_\nu}, a_{i_\mu} \in \Gamma; i_\nu, i_\mu = \overline{1, m}; \nu, \mu = \overline{1, l}).$$

Если в последовательности вместо $mng_{i_\mu}(a_{i_\mu})$ стоит специальный символ nil , то это значит, что значение признака a_{i_ν} не определено.

Примеры объектов, заданных явным набором признаков, многочисленны. В области медицинской диагностики в качестве таких объектов можно рассматривать, например, последовательность значений соответствующих признаков: частоту пульса, величину артериального давления, некоторые показатели анализа крови и т. д.

Перейдем к рассмотрению объектов, заданных неявным набором признаков. Пусть задано конечное множество объектов H , которое будем называть генеральной совокупностью объектов $X_i (X_i \in H)$, и пусть на этом множестве существует такое его разбиение на два подмножества $H1$ и $H2$, что $H1 \cap H2 = \emptyset$ и $H1 \cup H2 = H$. Из множества H выделим подмножество X (принцип выделения X из H в данной работе не рассматривается) и, назвав его выборкой, зададим на нем, как на множестве, такое его разбиение на два подмножества $X1$ и $X2$, что $X1 \cap X2 = \emptyset$; $X1 \cup X2 = X$; $X1 \subset H1$ и $X2 \subset H2$. (Приведенное разбиение множества X на подмножества $X1$ и $X2$ несколько идеализировано. Более реальным является следующее разбиение: $X1 \cap H1 \neq \emptyset$ и $X1 \cap H2 \neq \emptyset$, т. е. $X1 \not\subset H1$, при этом $X2 \subset H2$ или $X2 \cap H1 \neq \emptyset$ и $X2 \cap H2 \neq \emptyset$, но $X1 \cap X2 = \emptyset$).

Любой объект $x_\nu \in X$ характеризуется конечной последовательностью символов a_i из алфавита $\Gamma = \{a_1, a_2, \dots, a_m\}$, т. е.

$$x_\nu = a_{i_1} \rightarrow a_{i_2} \rightarrow \dots \rightarrow a_{i_l},$$

где $a_{i_j} \in \Gamma (i_j = \overline{1, m}; j = \overline{1, l} \text{ и } \exists_{x_j \in X} x_j (a_{i_\nu} = a_{i_\mu}) \text{ при } \nu \neq \mu)$.

Для описания объектов множества X в терминах «признак—значение» зададим конечное базовое множество признаков θ , значения которых определяются с помощью соответствующих им алгоритмов. В таблице приведено базовое множество признаков, используемое в конкретных системах СФП-1 и СФП-2, и указаны обозначения соответствующих им алгоритмов.

В данной статье не рассматривается вопрос о том, как были выбраны эти признаки и почему именно они составили базовое множество признаков. Заметим только, что было бы чрезвычайно заманчивым сформировать базовое множество из таких признаков, выделение значений которых человеком из объектов внешней среды осуществляется врожденными, встроенными в организм механизмами. Результаты, полученные нейрофизиологами и нейропсихологами [3, с. 147], позволяют думать о правомерности такой постановки вопроса.

В описании объектов рассматриваемого класса легко заметить по крайней мере одну особенность: для каждого из них значения признаков не заданы человеком заранее (как в предыдущем случае), а должны определяться самой системой. Если установление соответствия между каждым объектом и значениями первых двух признаков с помощью соответствующих алгоритмов A_1 и A_2 не вызывает затруднений, то процесс определения значений остальных признаков не так тривиален. В этом и состоит вторая особенность описания объектов, заданных неявным набором признаков.

Признаки			Алгоритм
обозначения	пояснения		
LNG	P_1	Количество символов в последовательности	A_1
SSN	P_2	Одинаковым символам входить в последовательность больше одного раза	A_2
SCS	P_3	Быть подмножеством заданного множества символов	A_3
SSS	P_4	Конечному множеству символов располагаться в определенной последовательности	A_4
DSS	P_5	Символам в последовательности находиться на определенном расстоянии по отношению друг к другу	A_5
NSL	P_6	Символам в последовательности слева направо находиться на определенных местах	A_6
NSR	P_7	Символам в последовательности справа налево находиться на определенных местах	A_7
OSB	P_8	Одному символу последовательности находиться на определенном расстоянии от начала последовательности	A_8
OSE	P_9	Одному символу последовательности находиться на определенном расстоянии от конца последовательности	A_9
EES	P_{10}	Равное число вхождений различных символов во все последовательности множеств X_1 или X_2	A_{10}
GES	P_{11}	Объединение числа вхождений различных символов во все последовательности множеств X_1 или X_2	A_{11}

В самом деле, если, например, считать, что объект характеризуется множеством n символов (порядок следования которых в данном случае не учитывается), то количество его подмножеств (значений 3-го признака) равно 2^n , а выбор конкретного значения данного признака для одного объекта не представляется возможным в связи с отсутствием критерия выбора того или иного значения. Поэтому значение таких признаков определяется путем анализа объектов множеств X_1 и X_2 . Для более строгого описания объектов рассматриваемого класса в терминах «признак—значение» введем ряд обозначений и определений. Тот факт, что каждому объекту x_v можно поставить в соответствие конечное множество значений p_i -признака, обозначим через $\{mng_j(p_i(x_v))\}$. Исключение составляет только первый признак, имеющий единственное значение для каждого объекта.

Определение 1. Под качественным признаком будем понимать такой признак, значение которого не отражает числовую характеристику объекта.

Например, качественными признаками будут порядковые номера, профессий (токарь — 1, врач — 2, тракторист—3 и т. д.). Для введенной нами системы признаков (см. таблицу) к качественным признакам относятся следующие: p_3, p_4 .

Определение 2. Будем говорить, что признак $p_i \in \theta$ имеет существенные значения, если выполняются два условия:

$$1) \forall x_v \in X_1 \exists p_i \in \theta (\cap \{mng_j(p_i(x_v))\} \neq \emptyset);$$

$$2) \forall x_v \in X_1 \exists p_i \in \theta \forall x_\mu \in X_2 ((\cap \{mng_j(p_i(x_v))\}) \cap (\cap \{mng_j(p_i(x_\mu))\})) = \emptyset),$$

где $\forall x$ — квантор всеобщности, а $\exists p$ — квантор существования.

Термин «существенный признак» означает, что признак имеет хотя бы одно существенное значение. Тот факт, что все объекты множества X_1 имеют существенный признак, обозначим $esnX_1$.

Определение 3. Будем говорить, что признак $p_{i_2} \in \theta$ зависит от признака $p_{i_1} \in \theta$, если определение значений признака p_{i_2} связано с одним из значений признака p_{i_1} .

Например, чтобы найти значения признака $DSS(p_5)$ для объекта x , (таблица), необходимо вначале определить значение признака $SSS(p_4)$. Для данного примера множество значений признака DSS запишется следующим образом:

$$\{mng_j(p_5(mng_i(p_4(x))))\}.$$

На рис. 1 в виде иерархической структуры представлено базовое множество признаков (таблица). Как видно из рисунка,

независимыми являются признаки *LNG*, *SCS*, *EES* и *GES*, а зависимыми — *SSN*, *SSS*, *OSB*, *OSE*, *DSS*, *NSL*, *NSR*.

Определение 4. Признаки называются логическими, если их значения указывают только на наличие или отсутствие этих признаков в анализируемом объекте.

Таким образом, логический признак может принимать только два значения: «истинно» или «ложно» («да» или «нет», «1» или «0» и т. д.). Легко видеть, что любой признак вместе со своим значением можно представить в виде логического признака. В самом деле, пусть значение признака *LNG*(p_1) равно 5, тогда на основании данного признака и его значения получим следующий логический признак, представленный в виде вопроса: «длина последовательности равна 5?».

Определение 5. Логический признак называется существенным, если он основывается на существенном признаке и его значении.

Обозначение *lgc(esnlX1)* будет говорить о том, что существенный признак объектов множества *X1* преобразован в логический существенный признак.

Определение 6. Под понятием будем понимать обобщенную информацию о множестве объектов, представленную в виде логической функции, связывающей значения существенных логических признаков. Через *сnpX1* будем обозначать понятие множества *X1* и представлять его как *lgc(esnlX1)*, т. е.

$$сnpX1 = lgc(esnlX1). \quad (1)$$

В ряде случаев целесообразнее сформировать *сnp X1*, используя *епр X2*. При этом

$$сnpX1 = \neg lgc(esnlX2). \quad (2)$$

Данные понятия представлены простейшими логическими функциями. Практически же при формировании понятий чаще всего приходится сталкиваться с таким множеством объектов *X*, что $\bigcap_{x, e \in X1} \bigcup_{p_i \in \Theta \nu} \{mng_{i\nu}(p_i(x))\} \neq \emptyset$ или

$$\bigcup_{x, e \in X1} \bigcap_{p_i \in \Theta \nu} \bigcup_{x_\mu \in X2} (\bigcap_{\nu} \{mng_{i\nu}(p_i(x_\nu))\}) \cap (\{mng_{i\mu}(p_i(x_\mu))\}) \neq \emptyset).$$

В этом случае *сnpX1* определяется уже более сложными логическими функциями:

$$сnpX1 = \bigvee_i lgc(esnlX1_i), X1_i \subset X1 \quad (3)$$

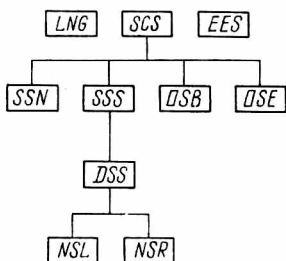


Рис. 1.

или

$$\text{снр}X1 = \bigcap_i (\forall \text{lgc}(\text{есnl}X2_i)), X2_i \subset X2, \quad (4)$$

$$\text{где } \bigcup_i X1_i = X1 \text{ и } \bigcup_i X2_i = X2.$$

В самом общем случае понятие, определенное на множестве объектов, заданных неявным набором признаков, будет описываться следующими выражениями:

$$\text{снр}X1 = (\forall \text{lgc}(\text{есnl}X1_i)) \wedge \bigcap_j (\forall \text{lgc}(\text{есnl}X2_j)), \quad (5)$$

$$\text{где } \bigcup_i X1_i = X1, \text{ а } \bigcup_j X2_j \subset X2, \text{ или}$$

$$\text{снр}X2 = \bigcap_i (\forall \text{lgc}(\text{есnl}X1_i)) \wedge (\forall \text{lgc}(\text{есnl}X2_j)), \quad (6)$$

$$\text{где } \bigcup_i X1_i \subset X1 \text{ и } \bigcup_j X2_j = X2.$$

Когда понятие $\text{снр}X1$ сформировано, то с его помощью можно определить принадлежность произвольного объекта $x_i (x_i \in H)$ множеству $X1$. Этот процесс будем обозначать через $\text{снр}X1(x_i)$. Если $\text{снр}X1(x_i) = 1$, то это значит, что $x_i \in H1$, а если $\text{снр}X1(x_i) = 0$, то $x_i \notin H1$. Первоначальная задача формирования понятий состоит в том, чтобы при $X1 \subset H1$ и $X2 \subset H2$ получить такой $\text{снр}X1$, что

$$\forall x_i \in H1 - X1 \quad \forall x_j \in H2 - X2 ((\text{снр}X1(x_i) = 1) \wedge (\text{снр}X1(x_j) = 0)). \quad (7)$$

Поскольку проблема выделения таких подмножеств $X1$ и $X2$ из соответствующих множеств $H1$ и $H2$, чтобы выполнялось условие (7), остается открытой, то в большинстве случаев рассматриваемая задача формирования понятий (см. выражение (7)) носит больше теоретический характер, чем практический. Для реальных же ситуаций более справедливо выражение (8):

$$\bigcap_{x_i \in H1 - X1} \bigcap_{x_j \in H2 - X2} ((\text{снр}X1(x_i) = 1) \wedge (\text{снр}X1(x_j) = 0)). \quad (8)$$

Система формирования понятий (СФП) — ошибаясь, «учиться на ошибках», — корректируя сформированные понятия, уменьшает процент ошибочных ответов. Сформированные понятия СФП хранит в виде «и/или» структур [2], что позволяет ей использовать приобретенные «знания» при формировании других понятий.

Построение такой СФП, которая каждое новое понятие формировала бы на основании своего «опыта», приобретенного в процессе формирования предыдущих понятий, и является одной из важнейших задач формирования понятий.

Перейдем теперь к рассмотрению блок-схемы алгоритма функционирования СФП-2 (рис. 2), разработанного в Харьковском институте радиоэлектроники. СФП-2 — это система индуктивного формирования понятий, представляемых в виде логических выражений (типа (1)–(6), а для СФП-1 — (1)–(2)). СФП-2 может работать с объектами, заданными неявным и явным набором признаков. Система реализована на алгоритмическом языке FORTRAN-IV.

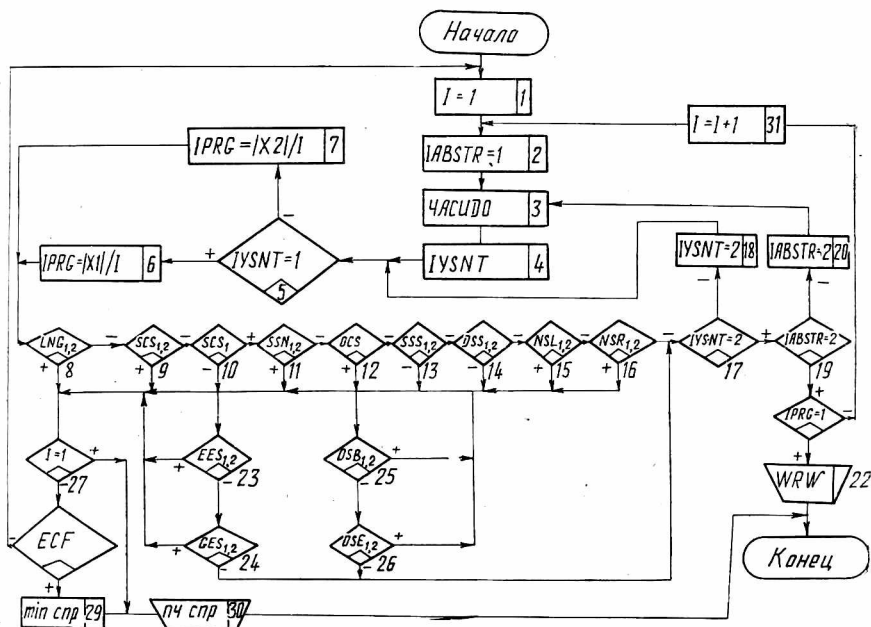


Рис. 2.

При описании блок-схемы алгоритма использовались следующие обозначения:

I — переменная, с помощью которой изменяются значения пороговых величин ($IPRG$) в соответствии с количеством объектов в подмножествах X_1 и X_2 (блоки 6 и 7);

$IABSTR$ используется для обозначения уровня абстракции. Если $IABSTR=2$, то СФП-2 переходит от конкретных значений символов ($IABSTR=1$) к более общим понятиям: русские гласные и согласные буквы, латинские гласные и согласные буквы, цифры и символы;

$|X_1|$ и $|X_2|$ обозначают мощность соответствующих подмножеств X_1 и X_2 ;

ЧАСИДО — блок частотного анализа символов и длин объектов;

$IYSNT$ — переменная, принимающая два значения. Если $IYSNT = 1$, то СФП-2 анализирует объекты подмножества X_1 , а если $IYSNT = 2$, то объекты подмножества X_2 .

Если при определении существенных значений тех или иных признаков проверяется выполнение только первого условия (см. определение 2), то это записывается как p_{i_1} , или например, для конкретного признака, как SCS_1 . Если же проверяется выполнение обоих условий, то вводится обозначение $p_{i_{1,2}}$ или $SCS_{1,2}$.

Система СФП-2 реализует следующий алгоритм формирования понятий (рис. 2).

1. Предпринимается попытка среди всех объектов множества X_1 на первом уровне абстракции отыскать существенное значение признака ($I=1$; $IABSTR=1$; $IYSNT=1$; $IPRG=|X_1|$. Блоки 1, 2, 4—6). Чтобы избежать полного перебора признаков, введен блок ЧАСИДО, который формирует алфавиты символов соответственно для объектов множеств X_1 и X_2 , подсчитывает частоту одноразового вхождения символов алфавитов в объекте соответствующих множеств X_1 и X_2 , а также частоту максимального вхождения каждого символа в объекты этих же множеств. Кроме этого, данный блок определяет длину каждого объекта как число символов, составляющих этот объект, и подсчитывает частоту встречаемости объектов определенной длины в соответствующих множествах X_1 и X_2 .

Анализ результатов работы блок ЧАСИДО позволяет из всего множества признаков выбрать подмножества таких признаков, значения которых теоретически могли бы быть существенными для всех объектов множества X_1 (блоки 8 (+или—), 9(+или—), 10 (+), 11—15 (+или—), 16 (+), 25 (+или—) и 26 (+). Если в результате анализа СФП-2 не может выделить такое подмножество признаков (блок 10 (—)), то она попытается найти значения признаков EES и GES (блоки 23, 24). Если существенное значение какого-либо признака найдено (блоки 8—9 (+), 11—16 (+), 23—26 (+)), то работа алгоритма оканчивается при $I = 1$ (блок 27 (+)).

2. В случае отрицательного исхода действий, выполняемых в п. 1, делается попытка выполнить аналогичную процедуру для объектов множества X_2 (блоки 16 (—), 24 (—), 26 (—), 17 (—), 18, 5, 7). Если удается найти существенное значение того или иного признака для всех объектов множества X_2 , то работа алгоритма оканчивается.

3. В случае неудачи СФП-2 переходит ко второму уровню абстракции (блоки 17 (+), 19 (—), 20, 3—4, 5 (+), 6) и все действия, описанные в пунктах 1 и 2, повторяются заново. В случае успеха алгоритм прекращает свою работу.

4. В противном случае (блок 19 (+)) СФП-2 переходит к изменению порогового значения (блоки 21 (—), 31, 2—4, 5 (+), 6), что приводит к попытке выделить во множествах X_1 или X_2 (в зависимости от значения $IYSNT$) такие подмножества мощ-

ностью не менее $IPRG$, все объекты которых содержали бы существенные значения какого-либо признака. Иными словами, после изменения порогового значения повторяются пункты 1—3. Так будет продолжаться до тех пор, пока понятие не будет сформировано (блоки 27 (—), 28 (+)). После минимизации понятия (блок 29) оно выводится на печать (блок 30). Сформированное понятие, в зависимости от сложности решаемой задачи, представляется в виде (1)—(6).

Если понятие не сформировано (блок 21 (+)), то это говорит о сбое в работе ЭЦВМ, при этом на печать выводится содержимое основных массивов и переменных СФП-2 (блок 22).

Рассмотренный алгоритм является сходящимся при формировании понятий любой сложности. Будем считать алгоритм формирования понятий *сходящимся*, если он обеспечивает формирование в конечное время понятия, разделяющего объекты обучающей выборки, принадлежащие X_1 или X_2 [1].

Теорема. *Время работы СФП-2 всегда конечно. Сформированное понятие полностью разделяет объекты обучающей выборки X на объекты, принадлежащие X_1 или X_2 .*

Справедливость данной теоремы очевидна, так как множества X_1 и X_2 конечны, сформированное понятие состоит из существенных значений признаков, охватывающих все объекты X_1 или X_2 , и СФП-2 всегда сформирует понятие, потому что при $IPRG=1$ СФП-2 будет работать с единичными подмножествами, для каждого объекта из которых существенное значение признака всегда будет найдено.

В заключение автор выражает глубокую благодарность профессору Э. Хангу за полезное обсуждение принципов построения СФП-2 и предложенного автором базового множества признаков.

СПИСОК ЛИТЕРАТУРЫ

1. Гладун В. П., Ващенко Н. Д. Методы формирования понятий на ЦВМ (обзор). — «Кибернетика», 1972, № 2, с. 107—112.
2. Нильсон Н. Искусственный интеллект. М., «Мир», 1973. 301 с.
3. Прибрам К. Языки мозга. М., «Прогресс», 1975. 464 с.

Поступила 20 декабря 1975 г.